Determining the Number of Communication Sources Using a Sensor Array

Eran Fishler[‡],Michael Grossman[†], and Hagit Messer[†]

[†] School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel, email: messer@eng.tau.ac.il
[‡] Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA, e-mail: eran.fishler@njit.edu

Abstract

Determining the number of sources in a received wave-field is a well known and a well investigated problem. In this problem, the number of sources impinging on an array of sensors is to be estimated. The common approach for solving this problem is to use an information theoretic criterion like the Minimum Description Length (MDL), or the Akaike Information Criterion. Under the assumption that the transmitted signals are Gaussian, the MDL estimator takes both simple and intuitive form. Therefore, this estimator is commonly used even when the signals known to be non-Gaussian communication signals. However, its ability to resolve signals (resolution capacity) is limited by the number of sensors, minus one. In this paper, we study the MDL estimator that is based the correct, non-Gaussian signal distribution of digital signals. We show that this approach leads to both improved performance and improved resolution capacity, that is - the number of signals that can be detected by the resulting MDL processor is larger than the number of array sensors. In addition, a novel asymptotic performance analysis, which can be used to predict the performance of the MDL estimator analytically, is presented. Simulation results support the theoretical conclusions.

I. Introduction

One of the fundamental problems in the field of array processing is to determine the number of signals that impinge on a passive sensors array. The problem arises in several fields of array processing such as radar, and seismology (see, among many others, [3], [5], [12], [13], [14], [17]). This question received considerable interest not only because of its nature, but also because most algorithms for direction-of-arrival estimation (hence DOA) assume a prior knowledge of the number of signals. Moreover, with the advent of new communication systems and the growing lack of frequency spectrum efficient techniques of communication are being developed. One of these approaches is to use antenna array to generate a narrow beam focused at the receiver. This, in-turn, enables to re-use the frequency spectrum for communication with other receivers. The method is being developed for cellular communication, where the users move and the cells should sense the advent of new users. New users are detected by monitoring the change in the number of users that the cell senses. The ability to detect the number of transmitters and their location is also required in spectrum monitoring and control applications.

The most common approach for estimating this number is to apply information theoretic criteria, like the Minimum Description Length (MDL) or the Akaike Information Criterion (AIC) [9]. Since 1985 [12], when a Gaussian version of the MDL was first suggested for estimating the number of narrow band sources impinging on an array of sensors, this MDL-based estimator became the standard tool for accomplishing this task.

A. Problem Formulation

Assume an array of p sensors and denote by $\mathbf{x}(t)$ the received, p-dimensional, signal vector at instance t. Denote by q the number of signals impinging on the array. A common model for the received signal vector is:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_q]$ is a $p \times q$ matrix composed of q p-dimensional vectors, $\mathbf{a}_1, \ldots, \mathbf{a}_q$. $\mathbf{A}$ is referred to as the steering matrix. $\mathbf{s}(t) = [s_1(t) \; \cdots \; s_q(t)]^T$ is a $q \times 1$ signals vector, and $\mathbf{n}(t)$ is a $p \times 1$ vector of the additive noise.

Many problems may be formulated using this simple, linear model (see [1], [2] and references therein). These problems differ by the structure of the mixing matrix, $\mathbf{A}$, by the assumed knowledge about the unknown parameters, or by the statistical modeling of both the signal and the noise. For example, $\mathbf{n}$ and/or $\mathbf{s}$ can be assumed Gaussian or non-Gaussian; the additive noise correlation matrix can be assumed white or not; or the mixing matrix, $\mathbf{A}$ can be assumed full rank or not.

In this paper we are concerned with the problem of estimating q when the sources are digital signals. Denote by $\mathcal{D} = [d_1, \ldots, d_{|\mathcal{D}|}]$ an arbitrary signal constellation, where $d_i \in \mathcal{C}$, and $|D|$ is the size of the constellation. In the sequel we make the following assumptions

1. The additive noise is zero mean, both spatially and temporal white, complex Gaussian random process, with correlation matrix $\sigma_n^2\mathbf{I}$.
2. The signal vector is uniformly distributed over $\mathcal{D}^q$, independent from snapshot to snapshot.
3. The additive noise and the signal vector are independent.

Based on these assumptions, the probability distribution function (pdf) of the received signal is

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_N | \mathbf{A}, \sigma_n^2) = \prod_{i=1}^{N} \frac{1}{|D|^q} \sum_{\mathbf{s} \in \mathcal{D}^q} \mathcal{CN}\left(\mathbf{x}_i - \mathbf{A}\mathbf{s}, \sigma_n^2\mathbf{I}\right) \tag{2}$$

where $\mathcal{CN}(\mathbf{x}, \mathbf{R}) = \frac{1}{\pi|\mathbf{R}|} e^{-\mathbf{x}^H\mathbf{R}^{-1}\mathbf{x}}$, and ^{H} denotes the complex transpose. We note that opposed to the common problem formulation [12], we do not assume that $\mathbf{A}$ is a full rank matrix. As will be shown in the sequel, the full rank requirement is not necessary for identifiably when the signals are digital signals.

Note that our model, (1), is also used to represent the reception of any multiple access system, e.g., code-division-multiple-access (CDMA) systems transmitting over flat fading channels [11]. Our problem is to estimate the number of sources, q, given N independent snapshots of the array output, $\mathbf{x}(t_1), \ldots, \mathbf{x}(t_N)$, described by the pdf of (2).

B. The MDL approach

The information theoretic criteria approach is a general one for choosing a model that fits the data mostly from a family of possible models [9], [15]. That is, given a parameterized family of probability densities, $f_{\mathbf{X}}\left(\mathbf{X}|\boldsymbol{\theta}^{(q)}\right)$, $\boldsymbol{\theta}^{(q)} \in \boldsymbol{\Theta}_q$ for various q, select $\hat{q}$ such that:

$$\hat{q} = \arg\min_q \left\{ -L\left(\hat{\boldsymbol{\theta}}^{(q)}\right) + P(q) \right\} \tag{3}$$

where $L\left(\boldsymbol{\theta}^{(q)}\right) = \log f_{\mathbf{X}}\left(\mathbf{X}|\boldsymbol{\theta}^{(q)}\right)$ is the log-likelihood of the measurements, denoted by $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$, $P(q)$ is a general penalty function associated with the q-th family, and $\hat{\boldsymbol{\theta}}^{(q)} = \arg\max_{\boldsymbol{\theta}^{(q)} \in \boldsymbol{\Theta}_q} \left\{ f_{\mathbf{X}}\left(\mathbf{X}|\boldsymbol{\theta}^{(q)}\right)\right\}$. $\hat{\boldsymbol{\theta}}^{(q)}$ is the maximum likelihood estimate of the unknown parameters given the q-th family of distributions.

The MDL estimator is a special case of (3) with a certain penalty function. It is given by minimizing the MDL metric [9], that is:

$$\hat{q}_{MDL} = \arg\min_q \mathrm{MDL(q)} = \arg\min_q \left\{ -\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + 0.5|\boldsymbol{\Theta}_q| \log(N) \right\} \tag{4}$$

where $\mathrm{MDL(q)} = \left\{ -\log\left(f\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + 0.5|\boldsymbol{\Theta}_q| \log(N) \right\}$, and $|\boldsymbol{\Theta}_q|$ is the number of free parameters in $\boldsymbol{\Theta}_q$. Since many problems are over parameterized, $|\boldsymbol{\Theta}_q|$ represents the minimum number of parameters that characterize $f\left(\mathbf{X}|\boldsymbol{\theta}^{(q)}\right)$ completely. It is well known that asymptotically, under certain regularity conditions, the MDL estimator minimizes the description length (measured in bits) of both the measurements, $\mathbf{X}$, and the model, $\hat{\boldsymbol{\theta}}^{(q)}$ [9].

Although in many problems associated with array processing, e.g., direction of arrival (DOA) estimation, one has some prior knowledge on the signals' statistical properties or on the array geometry, when the number of sources is estimated, this prior knowledge is usually ignored. The reason for this is that by assuming Gaussian sources and regardless of the array geometry, the resulting MDL estimator (4), termed the GMDL estimator, has a simple closed form expression given by

$$\hat{q}|_{GMDL} = \arg \min_{q=0,\ldots,p-1} \left\{ -N \log \frac{\prod_{i=q+1}^{p} l_i}{\left(\frac{1}{p-q}\sum_{i=q+1}^{p} l_i\right)^{p-q}} + \frac{1}{2}(q(2p-q)+1)\log N \right\} \tag{5}$$

where $l_1 \geq l_2 \geq \cdots \geq l_p$ are the eigenvalues of the empirical received signal's correlation matrix, $\hat{\mathbf{R}} = \frac{1}{N}\sum \mathbf{x}_i \mathbf{x}_i^H$. It is well known that the GMDL estimator is a consistent estimator of the number of sources [4].

In our problem the unknown parameters are: the matrix $\mathbf{A}$ and the noise level. Since we do not restrict our attention to full rank mixing matrices, when the number of sources is assumed to be q, the number of unknown parameters is $2pq+1$. Let us denote by $\hat{\mathbf{A}}_q$ and $\sigma^2_{n_q}$ the ML estimates of the unknown parameters assuming q sources. The MDL estimator (4) for our problem (2) becomes

$$\begin{aligned} \hat{q}|_{MDL} &= \arg \min_{q=0,1,\ldots} -\log\{f(\mathbf{x}_1,\ldots,\mathbf{x}_N|\hat{\mathbf{A}}_q,\sigma^2_{n_q}) + 0.5(2pq+1)\log N\} \\ &= \arg \min_{q=0,1,\ldots} \left\{ -\sum_{i=1}^{N} \log \frac{1}{|D|^q} \sum_{\mathbf{s}\in\mathcal{D}^q} \mathcal{CN}\left(\mathbf{x}_i - \hat{\mathbf{A}}_q \mathbf{s}, \hat{\sigma}^2_{n_q}\mathbf{I}\right) + 0.5(2pq+1)\log N \right\} \end{aligned} \tag{6}$$

The reader can already spot an important difference between the commonly used GMDL (5) and the exact MDL solution (6). While the GMDL can detect up to $p-1$ sources, the MDL estimator can detect any number of sources. This property of the MDL estimator will be discussed in section II.

C. Paper organization

The paper is organized as follows: section II discuss the resolution capacity of this problem, and the asymptotic characteristics of the MDL estimator. Section III is devoted for numerical study of the performance of the MDL estimator; whereas Section IV is devoted for analytical study of the performance of the MDL estimator. Section V provides a summary and some concluding remarks.

II. Identifiability and Consistency

A. Identifiability

Consider a parameterized family of probability density functions $f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. This family of densities is said to be *identifiable* if for every $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$, the divergence between $f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta})$ and $f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta}')$ is greater than zero, that is $D\left(f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta})\,||f_{\mathbf{X}}(\mathbf{x}|\boldsymbol{\theta}')\right) > 0$, where $D(f(\mathbf{x})||g(\mathbf{x})) = \int f \log \frac{f}{g} d\mathbf{x}$ is the

Kullback-Leibler divergence between $f(\mathbf{x})$ and $g(\mathbf{x})$. This condition insures that there is a one-to-one relationship between the parameter space and the statistical properties of the measurements.

The model order selection problem [10] discussed in Section I-A is unidentifiable if it is possible to find, for some $k \neq l$, two points in the parameter space, $\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_k$ and $\boldsymbol{\theta}_l \in \boldsymbol{\Theta}_l$ such that $f(\cdot|\boldsymbol{\theta}_k) = f(\cdot|\boldsymbol{\theta}_l)$. If this is the case, one cannot distinguish between k sources and l sources because the statistical distribution of the received measurements is insensitive to whether θ_k or θ_l was transmitted. The following Theorem establishes the identifiability of our problem.

Theorem 1: Let $\mathbf{A}$ and $\tilde{\mathbf{A}}$ be two arbitrary $p \times m$ and $p \times n$ matrices, respectively. In addition, let $\sigma_n^2 > 0$ and $\tilde{\sigma}_n^2 > 0$. The two probability distribution functions $f(\mathbf{x}|\mathbf{A}, \sigma_n^2)$ and $f(\mathbf{x}|\tilde{\mathbf{A}}, \tilde{\sigma}_n^2)$ are equal if and only if $m = n$, $\sigma_n^2 = \tilde{\sigma}_n^2$, and $\mathbf{A} = \tilde{\mathbf{A}}$ up to a permutations of the rows.

Proof of Theorem 1: See appendix A

Theorem 1 points to very important difference between the problem of detecting Gaussian sources and detecting digital sources. For the case of Gaussian sources, the problem is identifiable not only when the number of sources is smaller than the number of sensors but also when the mixing matrix is full rank. For the case of digital sources, on the other hand, the problem is identifiable for every number of sources and for every mixing matrix $\mathbf{A}$ whether full rank or not. This fact has a significant importance in communication systems since we can estimate the number of users utilizing a given channel (say using code division multiple access scheme) with only one antenna.

B. Consistency of the MDL Estimator

In the previous subsection it was proven that the estimation problem defined in section I-A is identifiable. Once the estimation problem has been shown to be identifiable, it is possible to infer the number of sources from the measurements. However for a specific estimator, the issue of consistency must be considered. In model order selection problems, the common performance measure is the probability of error, that is $\mathrm{P}_e = \mathrm{P}\{\hat{q} \neq q\} = 1 - \mathrm{P}_c$. In what follows the MDL estimator is proven to be a consistent estimator, that is $\lim_{N\to\infty} P_e = 0$.

Theorem 2: The MDL estimator is a consistent estimator of the number of sources.

Proof of Theorem 2: See Appendix B

Both the proof of Theorem 2 and its practical implications deserve special attention. The proof of Theorem 2 is divided into two parts. In the first part it is argued that asymptotically the probability of under-estimation, i.e., $\mathrm{P}\{\hat{q} < q\}$, approaches zero as $N \to \infty$; whereas in the second part it is argued that asymptotically the probability of over-estimation, i.e., $\mathrm{P}\{\hat{q} > q\}$, approaches zero as $N \to \infty$. The first part is easily proven based on a more general theorem which appears in [4]. The second part, however, is proven using the Wilks' Theorem. The reader might recall that the first published consistency proof of the GMDL was based on the Wilks' Theorem [12]. Later, it was demonstrated that the Wilks' Theorem can not be used in this problem [16], and an alternate proof, which is based on Taylor's expansion, was proposed in [15]. Since then, Taylor's expansion is the only tool used for these type of a proof. In what follows, we briefly explain why the Wilks' Theorem can be used in our proof.

Recall that for Gaussian signals it is required that the matrix $\mathbf{A}\mathbf{R}_\mathbf{S}\mathbf{A}^H$ is a rank q matrix. Therefore, assuming q sources the unknown parameters are: $[\lambda_1, \ldots, \lambda_q, \mathbf{v}_1, \ldots, \mathbf{v}_q, \sigma_n^2]$, where σ_n^2 is the unknown noise level, $\lambda_1 > 0, \ldots, \lambda_q > 0$ are the eigenvalues of the matrix $\mathbf{A}\mathbf{R}_\mathbf{S}\mathbf{A}^H$, and $\mathbf{v}_1, \ldots, \mathbf{v}_q$ are the eigenvectors corresponding to these eigenvalues. It is easily proven that the parameter space corresponds to q sources lies on the border of the parameter space that corresponds to $q+1$ sources, and not inside it. The intuition for this is very simple. The same way the point 0 is on the boarder of the parameter space $(0, T)$, the points $[\lambda_1, \ldots, \lambda_q, 0, \mathbf{v}_1, \ldots, \mathbf{v}_{q+1}, \sigma_n^2]$, which compose the parameter space corresponding to q sources, are on the border of the points $[\lambda_1, \ldots, \lambda_q, \lambda_{q+1}, \mathbf{v}_1, \ldots, \mathbf{v}_{q+1}, \sigma_n^2]$, which compose the parameter space corresponding to $q+1$ sources. The regularity conditions in the Wilks' Theorem requires that the parameter space corresponding to q sources be inside the parameter space corresponding to $q+1$ sources and not on its boarder. Since in the Gaussian case this condition does not hold, the Wilks' Theorem can not be used.

Consider the case of digital signals. From Theorem 1 it is clear that assuming q sources the parameter space is $[a_{11}, \ldots, a_{p1}, a_{21}, \ldots, a_{pq}, \sigma_n^2 > 0]$, where $a_{ij} \in \mathcal{C}$ for every $1 \leq i \leq p, 1 \leq j \leq q$. Now, the same way the point 0 is inside the parameter space $[-T, T]$, the points $[a_{11}, \ldots, a_{pq}, 0, \ldots, 0, \sigma_n^2 > 0]$, which compose the parameter space corresponding to q sources, are inside the points $[a_{11}, \ldots, a_{p(q+1)}, \sigma_n^2 > 0]$, which compose the parameter space corresponding to $q+1$ sources. Therefore, the Wilks' Theorem can be used in our problem for proving the the consistency of the MDL estimator.

Theorem 1 demonstrates that, at least theoretically, one can estimate the number of digital sources even when the number of receivers is one. The importance of Theorem 2 lies in demonstrating that indeed one can estimate the number of sources (transmitters) with diminishing probability of error even with one receiving antenna. This is of special importance in code division multiple systems (CDMA) that use blind multi-user detection. In these systems it is assumed that the number of users is smaller than the processing gain. However, the MDL estimator proposed in this paper can be used to estimate any number of users and hence can be used to increase the range of operation of blind multi-user detectors.

III. Simulation Study

In what follows we compare the performance of the GMDL estimator (5) and the MDL estimator (6), which exploits the special structure of the transmitted signals. We compare the performance of the two processors when either the number of sensors is larger than the number of sources or the number of sensors is smaller than the number of sources. Note that in the later the GMDL can not be used, and the MDL estimator is the only existing option. For simplicity we use BPSK signals, hence $\mathcal{D} = \{\pm 1\}$. In addition, we used the EM algorithm for computing the ML estimates of the unknown parameters [7].

First we consider the performance of the MDL and GMDL estimator in the presence of one source. Consider a uniform linear array with three sensors and one transmitter located at the array boresight. In the first simulation we study the performance of the GMDL and MDL estimators as a function of the number of snapshots, whereas in the second simulation we examine the performance of the GMDL and the MDL estimators as a function of the source signal to noise ratio (SNR). Figures 1 and 2 depict the probability of correct decision as a function of the number of snapshots and as a function of the source SNR, respectively. For the first simulation the SNR per element was set to $-6dB$, and for the second simulation the number of snapshots was set to 60. For each point in the graph 2000 Monte-Carlo runs were made.

In the graphs we see the clear advantage of the MDL over the GMDL estimator. The GMDL requires about 2dB additional power or about 3 times the number of snapshots in order to achieve the same performance as the MDL estimator that exploits the special signal structure. The uniform performance improvement of the MDL over the GMDL should not come as a surprise. In [6] it was demonstrated that by using additional *a priori* information one can improve the estimator performance considerably.

In the next set of simulations we examine the case of detecting more than one source. Figure 3 depicts the probability of correct detection of two sources as a function of their spatial separation. For the first scenario one source is located at angle $\theta = 0$ and the second at angle ρ where ρ is varied between $0°$ and $30°$. We consider two scenarios. In the first the sources' SNR was set to $-3dB$ and the number of snapshots was set to 50, while in the second the sources' SNR was set to $0dB$ and the number of snapshots was set to 100. Here again the MDL demonstrates uniform performance improvement over the GMDL. More surprisingly, however, is that with a sufficient SNR and or number of snapshots, the MDL can perfectly resolve two sources of no spatial separation. This is in agreement with Theorems 1 and 2 where it was shown that the MDL estimator can resolve any number of sources with only one receiver. The MDL estimator, therefore, uses the special signal structure and not spatial diversity to resolve multiple sources. Thus, the MDL estimator can resolve multiple sources even without any spatial separation between them, as figure 3 demonstrates.

As already mentioned, one of the main advantages of the MDL estimator is its ability to operate even in the presence of more sources than sensors. In the next and final simulation we explore this feature. We consider one receiving antenna and two equal power sources, transmitting with 0db SNR. Figure 4 depicts the probability of a correct decision as a function of the number of snapshots for the MDL estimator only (since the GMDL estimator is invalid in this scenario). It is evident from the graph that the MDL estimator is consistent estimator. Moreover, only limited number of snapshots are required for achieving reliable probability of detection.

IV. Performance Analysis of the MDL Estimator

The probability of error of the MDL estimator is an important figure of merit, and usually extensive simulations are carried out in order to study its performance. It is desired to have a simple tool for predicting the performance of the MDL estimator which can be used not only to study it but also to design a penalty function in (3) that meets some required performance. In what follows we provide such simple tool for predicting the performance of the MDL estimator.

Denote by P_M the probability of miss (the probability of under-estimation), *i.e.,* $\mathrm{P}_M = \mathrm{P}(\hat{q} < q)$, and by P_{FA} the probability of false alarm (the probability of over-estimation), *i.e.,* $\mathrm{P}_{FA} = \mathrm{P}(\hat{q} > q)$. The probability of error, denoted by P_e, is the sum of the probability of miss and the probability of false alarm, that is $\mathrm{P}_e = \mathrm{P}_M + \mathrm{P}_{FA}$. The probability of miss can be approximated by: $\mathrm{P}_M \approx$

$\mathrm{P}\left(\mathrm{MDL}(q-1) < \mathrm{MDL}(q)\right)$ while the probability of false alarm can be approximated by $\mathrm{P}_{FA} = \mathrm{P}\left(\mathrm{MDL}(q+1) < \mathrm{MDL}(q)\right)$. These approximations are due to the fact that the MDL function, MDL(q), is a convex function with a single minimum [4].

The following two lemmas establish asymptotic approximations for P_M and P_{FA}, respectively.

Lemma 1: Asymptotically, for large N

$$\mathrm{P}_{FA} \approx 1 - F_{\chi^2_{2p}}(2p \log N) \tag{7}$$

where $F_{\chi^2_{2p}}(\cdot)$ is cumulative distribution function of the chi-square random variable with $2p$ degrees of freedom

Proof of Lemma 1: See (20) in the proof of Theorem 2

Lemma 2: Asymptotically, for large N

$$\mathrm{P}_M \approx 1 - Q\left(\frac{-\mu}{\sigma}\right) \tag{8}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{\alpha^2}{2}} \mathrm{d}\alpha$ is the error function, $\mu = N\mathrm{E}\left(-\log f\left(\mathbf{x}|\mathbf{A}^*, \sigma_n^{2*}\right) + \log f\left(\mathbf{x}|\mathbf{A}, \sigma_n^2\right) + p \log N\right)$, and $\sigma^2 = NVar\left(-\log f\left(\mathbf{x}|\mathbf{A}^*, \sigma_n^{2*}\right) + \log f\left(\mathbf{x}|\mathbf{A}, \sigma_n^2\right)\right)$.
$\left(\mathbf{A}^*, \sigma_n^{2*}\right) = \arg\min_{[\mathbf{A}, \sigma_n^2] \in \mathbf{\Theta}_{q-1}} \{D\left(-\log f\left(\mathbf{x}|\mathbf{A}^*, \sigma_n^{2*}\right) || \log f\left(\mathbf{x}|\mathbf{A}, \sigma_n^2\right)\right)\}$, where $D(f||g)$ is the divergence between the two distributions.

Proof of Lemma 2: See Theorem 1 in [4]

These Lemmas provide a simple tool to approximate the performance of the MDL processor. The approximation improves with the number of snapshots or the SNR.

In order to validate our analysis we consider the following example. Assume an uniform linear array with three sensors and one source that transmits a BPSK signal at the array boresight. Figure 5 depicts the probability of detection $P_c = 1 - P_e$ which has been derived empirically by simulations, and the probability of detection predicted by the two Lemmas, as a function of the SNR. The number of snapshots was set to $N = 60$, and 2000 Monte-Carlo runs were made. From the figure one sees that even under non-asymptotic conditions ($N = 60$) our theoretical analysis predicts the probability of detection quit well, as the results of the theoretical analysis differ from the results of the simulation by no more than 1 dB. This demonstrates not only the validity of our analysis, but also its applicability as a synthesis tool.

V. Summary and Concluding Remarks

In this paper, we investigated the problem of estimating the number of communication sources impinging an array of sensors. We proved the resolution capacity is not limited by the number of sensors, and that one can estimate any number of sources. This is in contrast to the usual paradigm, which assumes that the number of resolvable sources is smaller than the number of sensors. We also proved that the MDL estimator is a consistent estimator of the number of sources.

The performance of the MDL estimator is shown to be uniformly superior over the commonly used GMDL estimator. In two important scenarios, namely, of more sources than sensors and of more than one source, the MDL estimator significantly outperforms the GMDL estimator. When more sources than sensors are to be detected, one can not use the GMDL estimator and the MDL is the only estimator that can solve the problem. When two or more sources exists, the MDL estimator exploits both the signal structure and the spatial separation between the sources, which results in substantial performance improvement over the GMDL estimator.

Due to the high complexity of the ML estimator, the complexity of the MDL estimator is very high. However, other almost-optimal algorithms, like the EM algorithm, can be used instead of the ML estimator. As the simulation results demonstrate, the use of such algorithms results in excellent performance. The low complexity of the EM algorithm makes the MDL estimator that uses the EM algorithm for estimating the unknown parameters very practical.

References

[1] Y. Bresler and A. Macovski. Exact Maximum Likelihood Parameter Estimation of Suprimosed Exponential Signals in Noise. *IEEE Trans. on Acustic Speech and Signal Processing*, ASSP-34(10):1081–1089, Oct. 1986.

[2] A. M. Bruckstein, T.-J. Shan, and T. Kailath. The Resolution of Overlapping Echos. *IEEE Trans. on Acustic Speech and Signal Processing*, ASSP-33(12):1357–1367, Dec. 1985.

[3] C. M. Cho and P. M. Djuric. Detection and Estimation of DOA's via Bayesian Predictive Densities. *IEEE Trans. on Signal Processing*, SP-42(11):3051–3060, Nov. 1994.

[4] E. Fishler, M. Grosmann, and H. Messer. Detection of Signals by Information Theoretic Criteria: General Asymptotic Performance Analysis. *IEEE Transactions on Signal Processing*, 50(5):1027–1036, May 2002.

[5] E Fishler and H. Messer. On the Use of Order Statistics for Improved Detection of Signals by the MDL Criterion. *IEEE Trans. on Signal Processing*, SP-48(8):2242–2247, Aug. 2000.

[6] E. Fishler and H. Messer. On the effect of a-priori information on performance of the MDL estimator. In *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2002.

[7] M. Grosman. Constelation Based Detection of the Number of Sources. Master's thesis, Tel Aviv University, Feb. 2001.

[8] E. Moulines, J.-F. Cardoso, and E. Gassiat. Maximum Likelihood for Blind Speration and Deconvolution of Noisy Signals using Mixture Models. In *International Conference of Acustic Speech and Signal Processing*, volume 5, pages 3617–3620, 1997.

[9] J. Rissanen. A Universal Prior for Integers and Estimation by Minimum Description Length. *Ann. Stat.*, 11(2):431–466, 1983.

[10] J. Rissanen. Universal Coding, Information Prediction, and Estimation. *IEEE Trans. on Information Theory*, IT-30(4):629–636, July 1984.

[11] S. Verdú. *Multiuser Detection*. Cambridge University Press, Cambridge, UK, 1998.

[12] M. Wax and T. Kailath. Detection of Signals by Information Theoretic Criteria. *IEEE Trans. on Acustic Speech and Signal Processing.*, ASSP-33(2):387–392, Feb. 1985.

[13] K. M. Wong, Q.-T Zhang, J. P. Reilly, and P. C. Yip. On Information Theoretic Criteria for Determining the Number of Signals in High Resolution Array Processing. *IEEE Trans. on Signal Processing*, SP-38(11):1959–1971, Nov. 1990.

[14] H. T. Wu, J. F. Yang, and F. K. Chen. Source Number Estimation Using Transformed Gerschgorin Radii. *IEEE Trans. on Signal Processing*, SP-43(6):1325–1333, Jun. 1995.

[15] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai. On Detection of the Number of Signals in the Presence of White Noise. *J. Multivariate Analysis*, 20(1):1–20, Jan. 1986.

[16] L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai. Remarks on Certain Criteria for Detection of Number of Signals. *IEEE Trans. Acoust., Speech, Signal Processing*, 35(2):129–132, Feb. 1987.

[17] A. M. Zoubir. Bootstrap Methods for Model Selection. *AEU-INTERNATIONAL JOURNAL OF ELECTRONICS AND COMMUNICATIONS*, 53:386–392, 1999.

Appendices

I. Proof of Theorem 1

In this appendix we prove that the problem is identifiable. In particular we prove that there is a one-to-one correspondent between the measurements' statistical distribution and the spatial scenario. We assume without loss of generality that BPSK signals are transmitted.

Recall that given the mixing matrix, $\mathbf{A}$, and the noise level, σ_n^2, the pdf of the received signal vector is

$$f(\mathbf{x}|\mathbf{A},\sigma_n^2) = \frac{1}{2^q} \sum_{\mathbf{s}\in\{\pm 1\}^q} \mathcal{CN}\left(\mathbf{x}-\mathbf{A}\mathbf{s},\sigma_n^2\mathbf{I}\right), \tag{9}$$

where we used the assumption that the transmitted signals are BPSK signals. Let $\mathbf{A}$ be a $p \times q$ complex matrix, and $\mathbf{x}$ a $p \times 1$ complex vector. From $\mathbf{A}$ we create the $2p \times q$ matrix denoted by

$\bar{\mathbf{A}}$, and from $\mathbf{x}$ we create the $2p \times 1$ vector denoted by $\bar{\mathbf{x}}$. $\bar{\mathbf{A}}$ and $\bar{\mathbf{x}}$ are defined as follows:

$$\bar{\mathbf{A}} = [\Re[\mathbf{A}]_{1:}\ \Im[\mathbf{A}]_{1:} \cdots \Re[\mathbf{A}]_{1:} \Im[\mathbf{A}]_{1:}] \tag{10}$$

$$\bar{\mathbf{x}} = [\Re[\mathbf{x}]_1,\ \Im[\mathbf{x}]_1, \ldots, \Re[\mathbf{x}]_p,\ \Im[\mathbf{x}]_p]^T \tag{11}$$

where $[\mathbf{A}]_{i:}$ denotes the ith row of the matrix $\mathbf{A}$, and $[\mathbf{x}]_i$ denotes the ith element of the vector $\mathbf{x}$.

It is easy to verify that,

$$\begin{aligned} f(\mathbf{x}|\mathbf{A}, \sigma_n^2) &= \frac{1}{2^q} \sum_{\mathbf{s} \in \{\pm 1\}^q} \mathcal{CN}\left(\mathbf{x} - \mathbf{A}\mathbf{s}, \sigma_n^2 \mathbf{I}\right) = \frac{1}{2^q} \sum_{\mathbf{s} \in \{\pm 1\}^q} \mathcal{N}\left(\bar{\mathbf{x}} - \bar{\mathbf{A}}\mathbf{s}, \sigma_n^2 \mathbf{I}\right) \\ &= \frac{1}{2^q} \sum_{\mathbf{s} \in \{\pm 1\}^q} \prod_{i=1}^{2p} \mathcal{N}\left([\bar{\mathbf{x}}]_i - [\bar{\mathbf{A}}]_{i:}\mathbf{s}, \sigma_n^2\right) \triangleq f(\bar{\mathbf{x}}|\bar{\mathbf{A}}, \sigma_n^2). \end{aligned} \tag{12}$$

where $\mathcal{N}\left(x, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$. Furthermore, $f(\bar{\mathbf{x}}|\bar{\mathbf{A}}, \sigma_n^2)$ can be written as follows:

$$\begin{aligned} & f(\bar{\mathbf{x}}|\bar{\mathbf{A}}, \sigma_n^2) = \frac{1}{2^q} \sum_{\mathbf{s} \in \{\pm 1\}^q} \prod_{i=1}^{2p} \mathcal{N}\left([\bar{\mathbf{x}}]_i - [\bar{\mathbf{A}}]_{i:}\mathbf{s}, \sigma_n^2\right) \\ & = \prod_{i=1}^{2p} \mathcal{N}\left([\bar{\mathbf{x}}]_i, \sigma_n^2\right) * \frac{1}{2}\left(\prod_{i=1}^{2p} \delta([\bar{\mathbf{x}}]_i - [\bar{\mathbf{A}}]_{i,1}) + \prod_{i=1}^{2p} \delta([\bar{\mathbf{x}}]_i + [\bar{\mathbf{A}}]_{i,1})\right) \\ & * \frac{1}{2}\left(\prod_{i=1}^{2p} \delta([\bar{\mathbf{x}}]_i - [\bar{\mathbf{A}}]_{i,2}) + \prod_{i=1}^{2p} \delta([\bar{\mathbf{x}}]_i + [\bar{\mathbf{A}}]_{i,2})\right) * \cdots * \frac{1}{2}\left(\prod_{i=1}^{2p} \delta([\bar{\mathbf{x}}]_i - [\bar{\mathbf{A}}]_{i,q}) + \prod_{i=1}^{2p} \delta([\bar{\mathbf{x}}]_i + [\bar{\mathbf{A}}]_{i,q})\right) \end{aligned} \tag{13}$$

Now, the moment generating function of $\mathbf{x}$ given $\mathbf{A}$ and σ_n^2, or equivalently of $\bar{\mathbf{x}}$ given $\bar{\mathbf{A}}$ and σ_n^2 is

$$\phi_{\bar{\mathbf{x}}}(\omega|\bar{\mathbf{A}}, \sigma_n^2) = E\left\{e^{-j\boldsymbol{\omega}^T \bar{\mathbf{x}}}\right\} = \frac{1}{2^q} \prod_{i=1}^{2p} e^{\sigma_n^2 ||\boldsymbol{\omega}||^2} \prod_{i=1}^{q} \left[e^{-j\boldsymbol{\omega}^T [\bar{\mathbf{A}}]_{:i}} + e^{-j\boldsymbol{\omega}^T [\bar{\mathbf{A}}]_{:i}}\right] \tag{14}$$

We now turn to prove the theorem. It is easily seen that if $m = n$, $\sigma_n^2 = \tilde{\sigma}_n^2$, and $\mathbf{A}$ is equal to $\tilde{\mathbf{A}}$ up to column permutation, $f(\mathbf{x}|\mathbf{A}, \sigma) = f(\mathbf{x}|\tilde{\mathbf{A}}, \tilde{\sigma}_n^2)$. This proves the sufficiency.

We now turn to prove the necessity. We first prove that it is necessary that $\sigma_n^2 = \tilde{\sigma}_n^2$.

Lemma 3: Let $\mathbf{A}$ and $\tilde{\mathbf{A}}$ be two arbitrary $p \times m$ and $p \times n$ matrices, respectively. In addition, let $\sigma_n^2 > 0$ and $\tilde{\sigma}_n^2 > 0$. If $f(\mathbf{x}|\mathbf{A}, \sigma_n^2) = f(\mathbf{x}|\tilde{\mathbf{A}}, \tilde{\sigma}_n^2)$, then $\sigma_n^2 = \tilde{\sigma}_n^2$.

Proof of Lemma 3: Assume $f(\mathbf{x}|\mathbf{A}, \sigma_n^2) = f(\mathbf{x}|\tilde{\mathbf{A}}, \tilde{\sigma}_n^2)$. Consequently,

$$\begin{aligned} \phi_{\bar{\mathbf{x}}}(\omega|\bar{\mathbf{A}}, \sigma_n^2) &= \frac{1}{2^m} \prod_{i=1}^{2p} e^{\sigma_n^2 ||\boldsymbol{\omega}||^2} \prod_{i=1}^{m} \left[e^{-j\boldsymbol{\omega}[\bar{\mathbf{A}}]_{:i}} + e^{-j\boldsymbol{\omega}[\bar{\mathbf{A}}]_{:i}}\right] \\ &= \frac{1}{2^n} \prod_{i=1}^{2p} e^{\tilde{\sigma}_n^2 ||\boldsymbol{\omega}||^2} \prod_{i=1}^{n} \left[e^{-j\boldsymbol{\omega}[\bar{\tilde{\mathbf{A}}}]_{:i}} + e^{-j\boldsymbol{\omega}[\bar{\tilde{\mathbf{A}}}]_{:i}}\right] = \phi_{\bar{\mathbf{x}}}(\omega|\bar{\tilde{\mathbf{A}}}, \tilde{\sigma}_n^2). \end{aligned} \tag{15}$$

Assume that $\sigma_n^2 \neq \tilde{\sigma}_n^2$, and assume without loss of generality that $\sigma_n^2 > \tilde{\sigma}_n^2$. Dividing both sides of (15) by $\prod_{i=1}^{2p} e^{\tilde{\sigma}_n^2 ||\boldsymbol{\omega}||^2}$ results in

$$\frac{1}{2^m}\prod_{i=1}^{2p} e^{\left(\sigma_n^2-\tilde{\sigma}_n^2\right)||\boldsymbol{\omega}||^2}\prod_{i=1}^{m}\left[e^{-j\boldsymbol{\omega}[\bar{\mathbf{A}}]_{:i}}+e^{-j\boldsymbol{\omega}[\bar{\mathbf{A}}]_{:i}}\right]=\frac{1}{2^n}\prod_{i=1}^{n}\left[e^{-j\boldsymbol{\omega}[\bar{\tilde{\mathbf{A}}}]_{:i}}+e^{-j\boldsymbol{\omega}[\bar{\tilde{\mathbf{A}}}]_{:i}}\right]. \tag{16}$$

Taking the inverse Fourier transform the left and right hand side of (16) results in

$$\begin{aligned}
&\prod_{i=1}^{2p}\mathcal{N}\left([\bar{\mathbf{x}}]_i,\sigma_n^2-\tilde{\sigma}_n^2\right) * \frac{1}{2}\left(\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i-[\bar{\mathbf{A}}]_{i,1})+\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i+[\bar{\mathbf{A}}]_{i,1})\right)\\
&*\frac{1}{2}\left(\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i-[\bar{\mathbf{A}}]_{i,2})+\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i+[\bar{\mathbf{A}}]_{i,2})\right)*\cdots\\
&*\frac{1}{2}\left(\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i-[\bar{\mathbf{A}}]_{i,m})+\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i+[\bar{\mathbf{A}}]_{i,m})\right)=\frac{1}{2}\left(\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i-[\bar{\tilde{\mathbf{A}}}]_{i,1})+\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i+[\bar{\tilde{\mathbf{A}}}]_{i,1})\right)\\
&*\frac{1}{2}\left(\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i-[\bar{\tilde{\mathbf{A}}}]_{i,2})+\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i+[\bar{\tilde{\mathbf{A}}}]_{i,2})\right)*\cdots\\
&*\frac{1}{2}\left(\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i-[\bar{\tilde{\mathbf{A}}}]_{i,n})+\prod_{i=1}^{2p}\delta([\bar{\mathbf{x}}]_i+[\bar{\tilde{\mathbf{A}}}]_{i,n})\right).
\end{aligned} \tag{17}$$

It is now obvious that $\sigma_n^2 = \tilde{\sigma}_n^2$ is necessary condition for the equivalence between the two sides of equation (17). □

We now proceed to prove that $m = n$ is also necessary condition for $f(\mathbf{x}|\mathbf{A},\sigma_n^2) = f(\mathbf{x}|\tilde{\mathbf{A}},\tilde{\sigma}_n^2)$. Since $\sigma_n^2 = \tilde{\sigma}_n^2$, (17) becomes

$$\frac{1}{2^m}\sum_{\mathbf{s}\in\{\pm 1\}^m}\delta(\bar{\mathbf{x}}-\bar{\mathbf{A}}\mathbf{s})=\frac{1}{2^n}\sum_{\mathbf{s}\in\{\pm 1\}^n}\delta(\bar{\mathbf{x}}-\bar{\tilde{\mathbf{A}}}\mathbf{s}) \tag{18}$$

Consider the left hand side of (18), and assume without loss of generality that the first row of $\bar{\mathbf{A}}$ does not contain any zeros. In this case the maximum value of $[\bar{\mathbf{x}}]_1$ is $\sum_{j=1}^{q}|[\bar{\mathbf{A}}]_{1j}| \triangleq \alpha$. Consequently, the coefficient of product that includes $\delta([\bar{\mathbf{x}}]_1-\alpha)$ is $\frac{1}{2^m}$.

Now assume that parts of the elements of the first row. Since the first and second row of $\bar{\mathbf{A}}$ are related to the same source, if $[\bar{\mathbf{A}}]_{1i} = 0$ then $[\bar{\mathbf{A}}]_{2i} \neq 0$. Let $G = \{i|[\bar{\mathbf{A}}]_{1i} \neq 0\}$, $\bar{G} = \{i|[\bar{\mathbf{A}}]_1 i \neq 0\}$, and $\mathbf{S} = \{\mathbf{s}|\mathbf{s}\in\{\pm 1\}^q[\mathbf{s}]_i = \frac{[\bar{\mathbf{A}}]_{1i}}{|[\bar{\mathbf{A}}]_{1i}|}\ i \in G\}$. Denote by $|G|$ the number of elements in G, . It is easily seen that $|\mathbf{S}| = 2^{|\bar{G}|}$. Since $[\bar{\mathbf{A}}]_{2i} \neq 0 \forall i \in \bar{G}$, there exists one element in $\mathbf{S}$ denoted by $\tilde{\mathbf{s}}$ such that $[\bar{\mathbf{A}}]_{2:}\tilde{\mathbf{s}}$ is maximized. Therefore again, coefficient of product that includes $\delta([\bar{\mathbf{x}}]_1-\bar{\mathbf{A}}\tilde{\mathbf{s}})$ is $\frac{1}{2^m}$.

Using the same reasoning, we can demonstrate that there it least the coefficient of one product in the right hand side of (18) is 2^n. Therefore, it is necessary that $n = m$ for the equality to hold.

The uniqueness of $\bar{\mathbf{A}}$ is now easily proven from the (18) by noting that the delta functions having the coefficients have to be on same point in R^{2p}.

□

II. Proof of Theorem 2

In this appendix the consistency of the MDL estimator is proven. Specifically it is shown that the probability of error of the MDL estimator converges to zero as the number of snapshots increases to infinity. An error event will occur if and only if there exists $k \neq q$ such that $\mathrm{MDL}(q) - \mathrm{MDL}(k) > 0$. Thus in order to prove the lemma it suffice to prove that for every $k \neq q$, $\mathrm{P}\left(\mathrm{MDL}(q) - \mathrm{MDL}(k) > 0\right) \to 0$.

Assume $k < q$. It was previously shown that under very weak conditions the probability of miss of every MDL estimator converges to zero as $N \to \infty$ [4]. In particular the probability of miss of the MDL estimator considered in this paper, which satisfies the condition stated in [4], converges to zero as $N \to \infty$.

Now, assume that $k > q$. $\mathrm{MDL}(q) - \mathrm{MDL}(k)$ can be written as follows:

$$\begin{aligned}\mathrm{MDL}(q) - \mathrm{MDL}(k) &= -\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + 0.5(2pq+1)\log N + \log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(k)}\right)\right) - 0.5(2pk+1)\log N \\ &= -\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + \log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(k)}\right)\right) + 0.5(2p(q-k))\log N. \qquad (19)\end{aligned}$$

In what follows we are going to use the Wilks' Theorem to compute the asymptotic distribution of $\mathrm{MDL}(q) - \mathrm{MDL}(k)$.

Consider the difference between the log-likelihoods given q and $k > q$ sources, that is $-2\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + 2\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(k)}\right)\right)$. According to the Wilks' Theorem, under some regularity conditions, asymptotically $-\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + \log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(k)}\right)\right)$ is distributed as a chi-square random variable with $2p(k-q)$ degrees of freedom. The regularity conditions are the ones that insure that the MLE estimate is unique, efficient, and asymptotically normal. In [8], it shown that these regularity conditions hold in our problem.

Using the asymptotic distribution of $-\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + \log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(k)}\right)\right)$, we can compute the asymptotic probability of the event $\{\mathrm{MDL}(q) - \mathrm{MDL}(k) > 0\}$.

$$\begin{aligned} \mathrm{P}\left(\mathrm{MDL}(q) - \mathrm{MDL}(k) > 0\right) &= \mathrm{P}\left(-\log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(q)}\right)\right) + \log\left(f_{\mathbf{X}}\left(\mathbf{X}|\hat{\boldsymbol{\theta}}^{(k)}\right)\right) > 0.5(2p(q-k))\log N\right) \\ &\overset{N\to\infty}{\to} \mathrm{P}\left(\chi^2_{2p(k-q)} > (2p(q-k))\log N\right) \overset{N\to\infty}{\to} 0. \end{aligned} \tag{20}$$

We have proven that for $k \neq q$, the probability of the event $\mathrm{P}\left(\mathrm{MDL}(q) - \mathrm{MDL}(k) > 0\right) \overset{N\to\infty}{\to} 0$. Therefore, the probability of error approaches to zero as well. $\square$

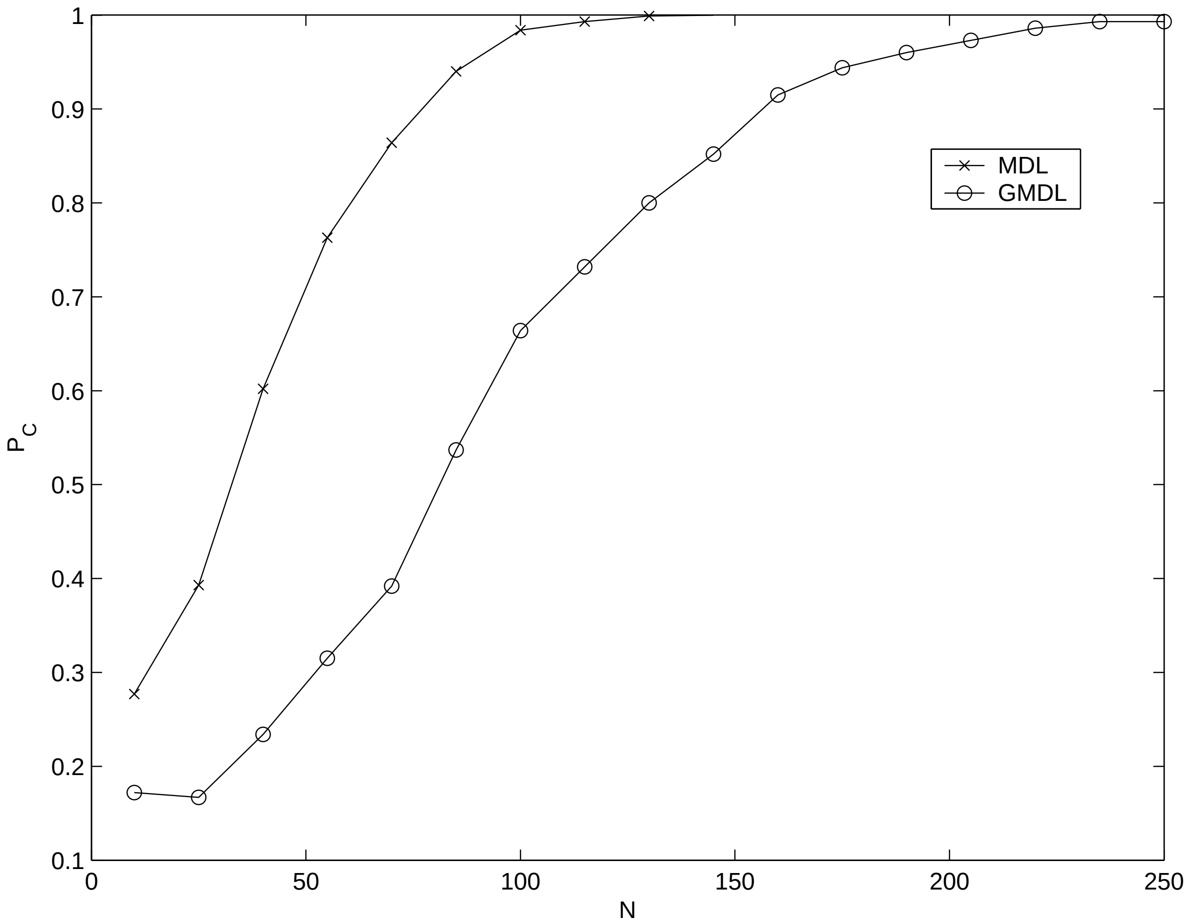

Fig. 1. The probability of a correct decision of the GMDL and MDL estimators for a single source as a function of the number of snapshots, N. ULA with $p = 3$, $\theta = 0$, $SNR = -6dB$

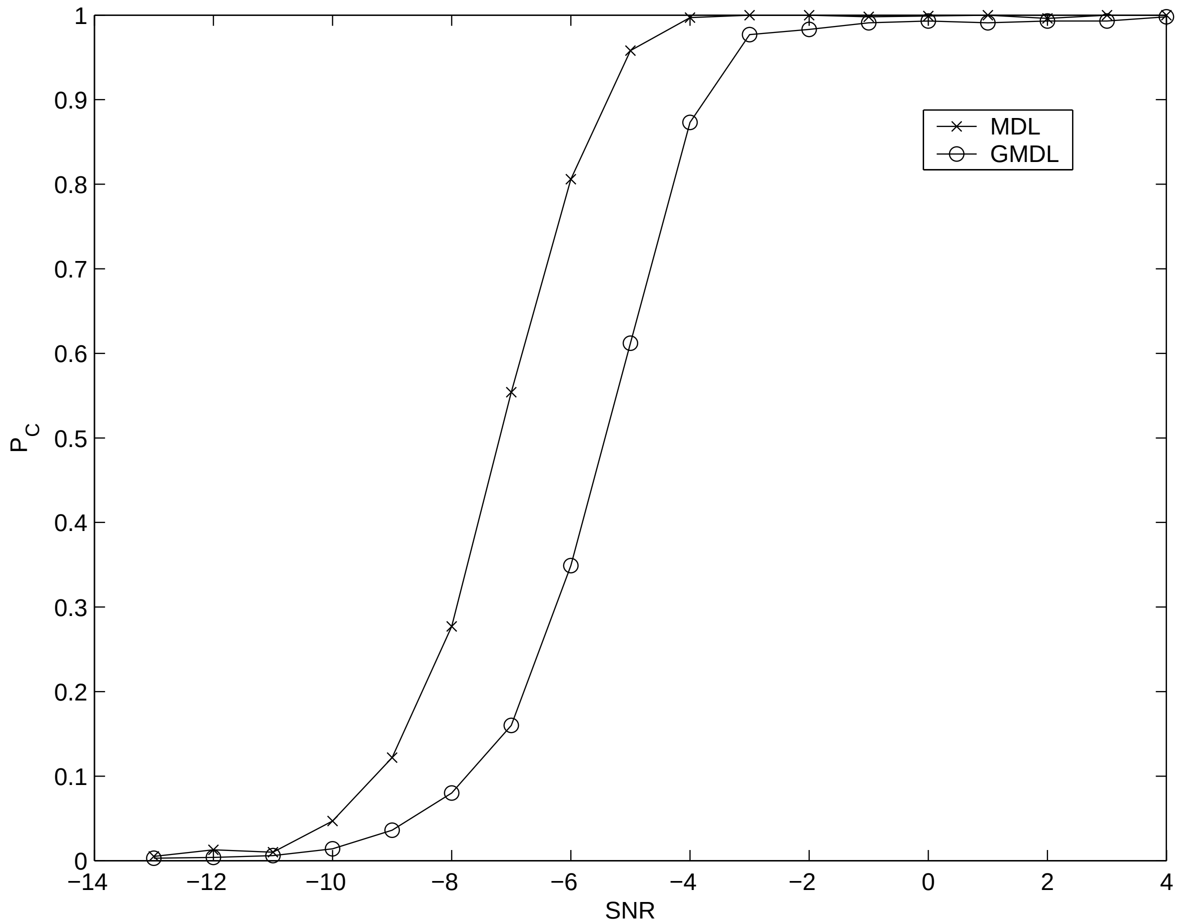

Fig. 2. The probability of a correct decision of the GMDL and MDL estimators for a single source as a function of the SNR. ULA with $p = 3$, $\theta = 0$, $N = 60$

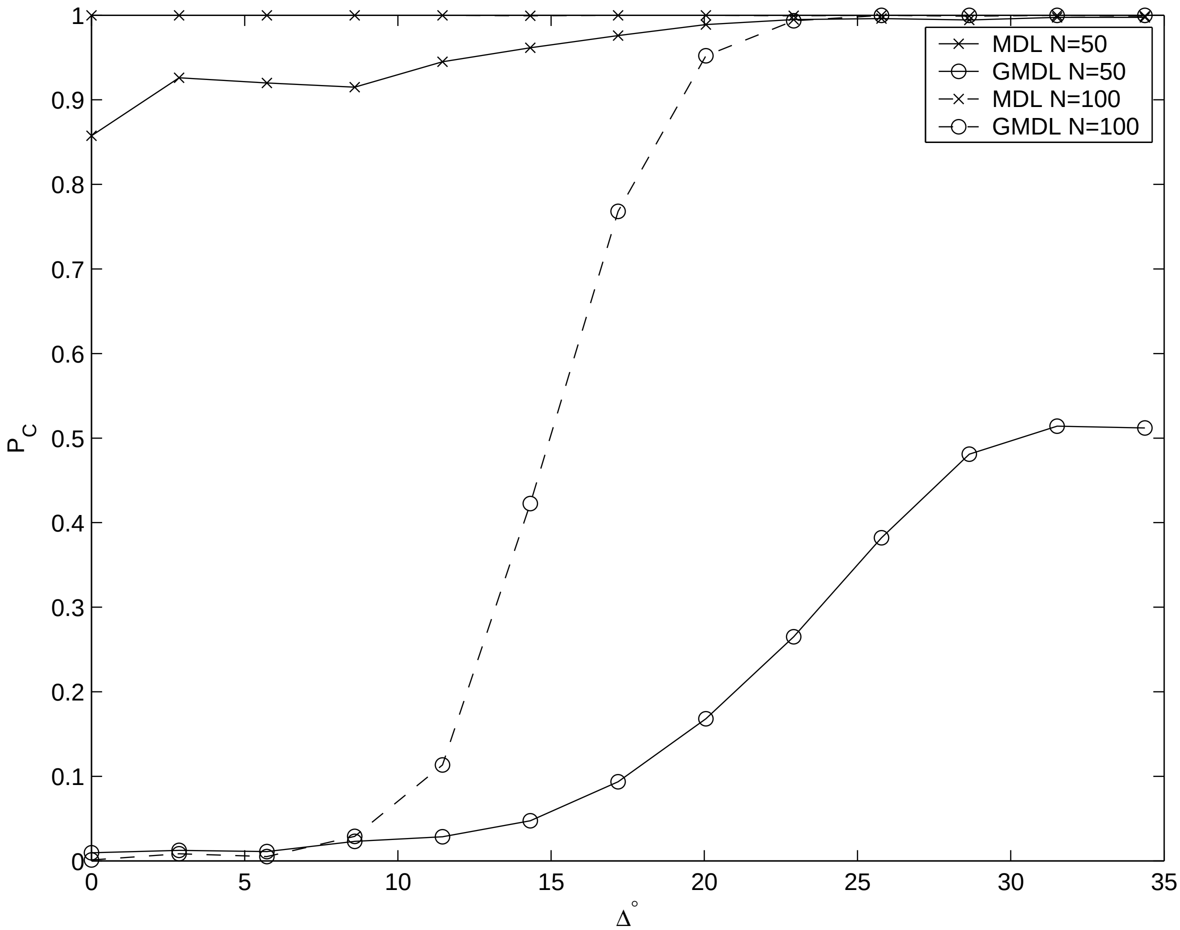

Fig. 3. The probability of a correct decision of the GMDL and MDL estimators for two equal-power sources as a function of the source separation. ULA with $p = 3$, $\theta_1 = 0°$, $N = 50$, $SNR = -3dB$, and $N = 100$, $SNR = 0dB$

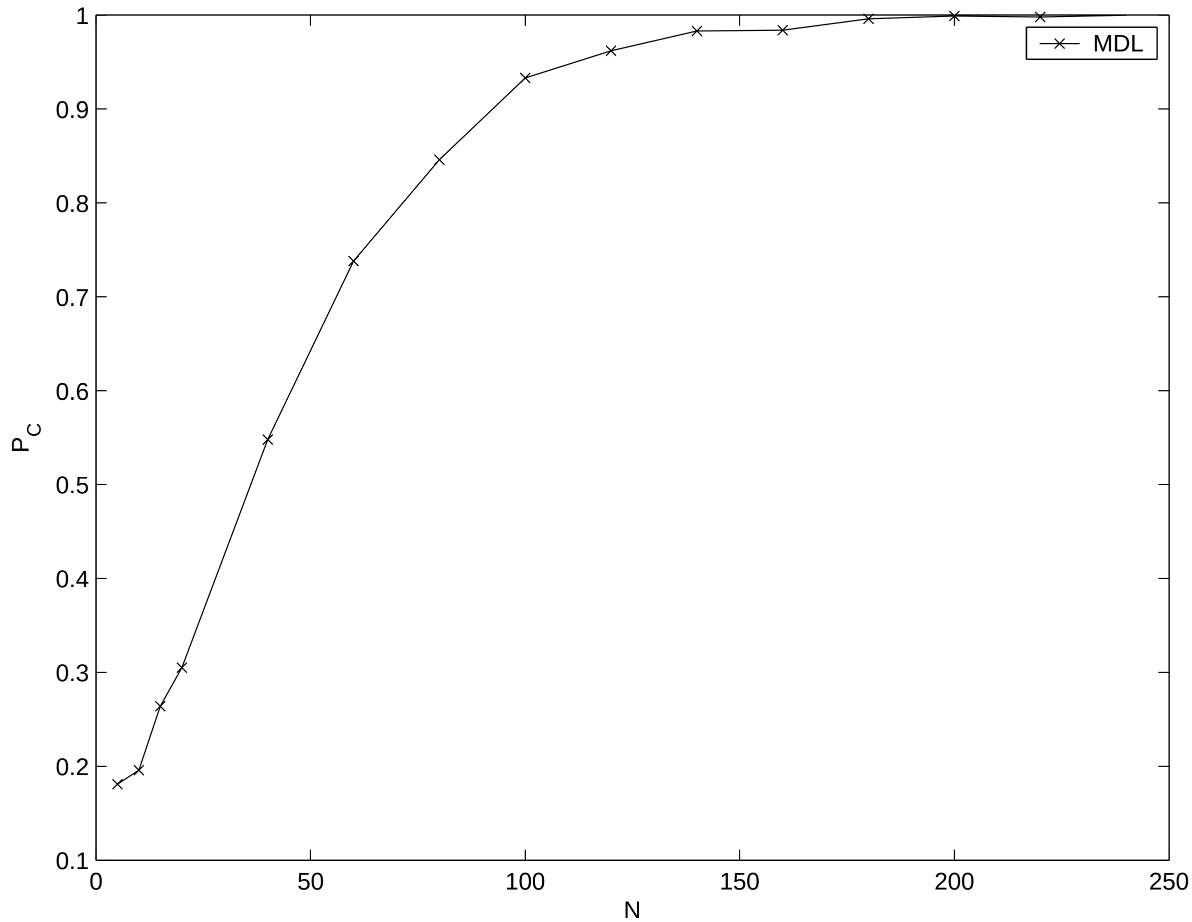

Fig. 4. The probability of a correct decision of the MDL estimators for two equal-power sources as a function of snapshots, using one receiver only and two equal power transmitters. $\theta_1 = 0$, $\theta_2 = 5^\circ$, $SNR = 0dB$

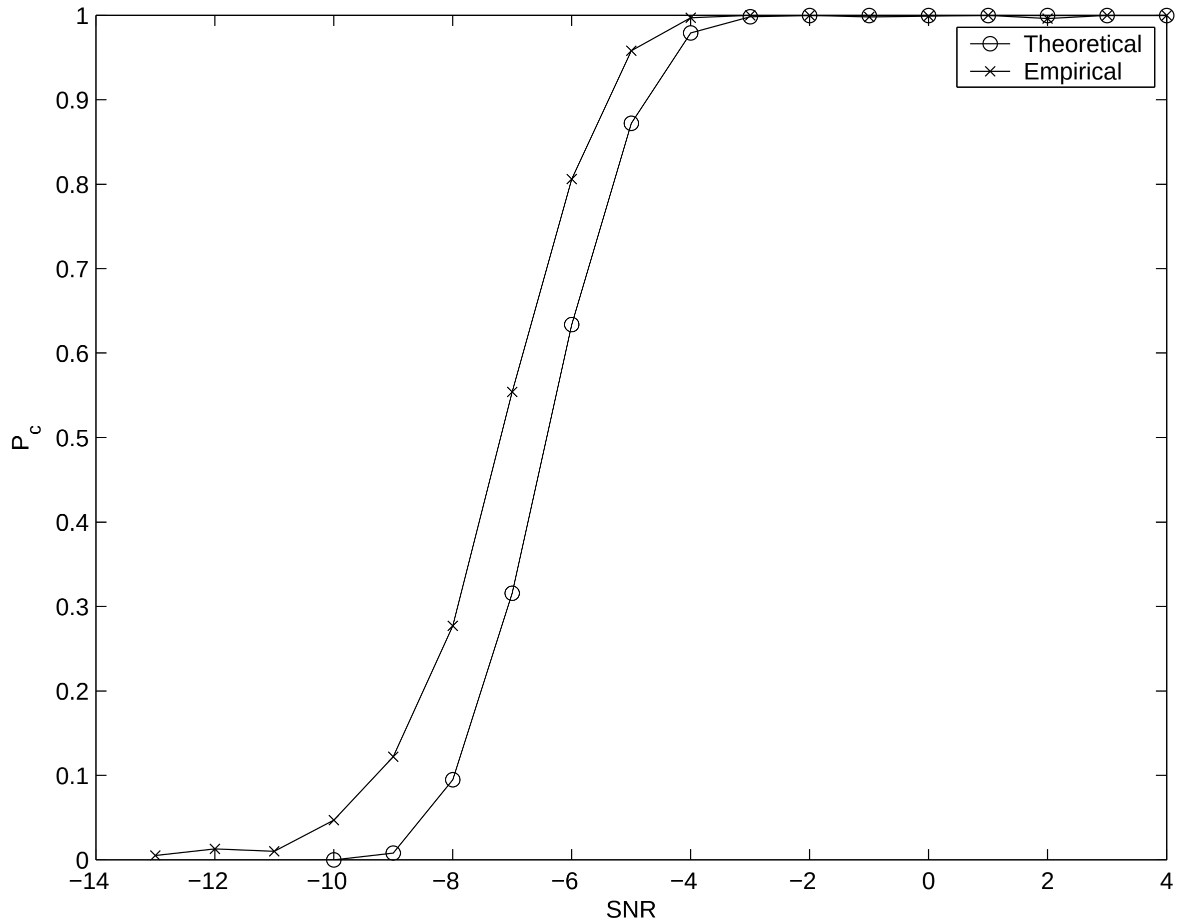

Fig. 5. The probability of a correct decision of the MDL estimators for a single source as a function the SNR. Empirical vs Theoretical results for ULA with p=3,$\theta = 0$, $N = 60$.